Persistent, mobile, or volatile? Long-run trends of earnings dynamics in Italy

Francesca Subioli (Sapienza University of Rome)

Michele Raitano (Sapienza University of Rome)

Provisional draft. Please do not quote
This version 16 June 2021

Abstract

The aim of this article is to describe the long-run patterns of earnings inequality in Italy since the 1970s, distinguishing persistent inequality, earnings mobility and volatility, and to study the relationship between mobility and volatility. Using a longitudinal dataset built matching survey and administrative data, and focusing on men employed in the private sector, we apply various methodologies to compare earnings dynamics experienced in the mid of working career – i.e., during ages from 35 to 45 – by 7 cohorts of workers born in each 5-year period from 1940 to 1974. We find that the general rise in total within-cohort earnings inequality for male workers is coupled with a steep rising contribution of permanent, between-workers inequality, and that the 'within-worker' inequality component is almost equally divided into mobility and volatility. Moreover, we find that experiencing a positive income trend is associated with lower volatility, and that the steepness of the income trend is positively related to volatility only for downward earnings trajectories.

JEL Codes: D31

Key words: Earnings inequality; Income mobility; Income volatility; Italy

1. Introduction

In recent years, the economic literature and the economic policy debate have been increasingly concerned with the rise in earnings and income inequality experienced since the last decades of the 20th century in most high-income countries (see, among others, OECD 2008 and 2011). Despite inequality is shaped by complex processes acting through various mechanisms and is influenced by several possible determinants (Atkinson 2015), a shared wisdom argues that these trends have been mainly due to processes acting in the markets and, specifically, the bulk of the increase in inequality seems attributable to the rise in earnings dispersion (Salverda et al. 2014).

However, also because of the lack of proper longitudinal data, most analyses about trends in earnings inequality provide pictures of what happened at various points in time (i.e., years) without tracking the same individual over time. In other terms, they usually focus on cross-sectional inequality neglecting intragenerational inequality and intertemporal mobility (Burkhauser and Couch 2009). Nevertheless, independently of the size of annual inequality, observing individual income dynamics is crucial to assess the characteristics of the process shaping inequality and its consequences on macroeconomic outcomes and individual wellbeing. As pointed out by OECD (2018), a society with a certain level of income inequality where individuals change their positions in the income ladder year by year faces different challenges with respect to a society with the same (or also a lower) level of inequality but where individuals never change their positions during their life.

The growing availability of long panel data, especially provided from administrative sources, has increasingly allowed researchers to investigate what happens over time to each worker, who may deviate (more or less frequently, more or less far away) from his/her average income throughout the career. The knowledge of cross-sectional earnings inequality can therefore be enriched by looking dynamically at how workers belonging to the same generation experience income changes, both with respect to their own starting position and their permanent income, and in comparison the other workers' situation. The idea is that, since in general individuals are concerned not only with the average income receipts over a certain period, but also with their pattern over time, and they prefer a stable stream to a fluctuating one (Shorrocks 1978), having a stable stream of income may be considered welfare-enhancing per se. If this is the case, the policy concern should deal as much with the level of income as with its dynamics over the career.

As discussed in Section 2, no univocal concepts, methods, and measures have emerged so far in the economic literature dealing with dynamic aspects of inequality (D'Ambrosio 2018). In this article, we focus on the concept of income mobility, which, following OECD (2018), can be roughly defined as the movements – downward or upward – of a given individual (independently of his or her starting point) through the distribution of income over time, by either relating one's current income to one's past income levels – which is referred to as absolute income mobility – or relating one's current income relative position in the income distribution to one's past position – which will be referred to as positional (or relative) income mobility.

To observe long-run patterns of earnings dynamics, we focus on Italy – a country characterised by a steep rise in labour income inequality in the last decades driven by rising gaps in both employment opportunities (also driven by a long-lasting process of labour market deregulation; Struffolino and Raitano 2020) and wages among individuals (Franzini and Raitano 2019). In Section 3 we provide an overall description of the structural changes in the Italian labour market from the 1970s. Making use of an extremely rich longitudinal dataset built matching survey and administrative longitudinal data (see Section 4), we contribute to the literature on income mobility from two perspectives: (i) we disentangle income patterns of male employees in the private sector followed in their mid-age (from age 35 to 45), to distinguish persistent inequality from earnings mobility and short-term volatility; (ii) we compare earnings trajectories of individuals belonging to several birth cohorts – i.e. workers born in each 5-year period from 1940 to 1974 – thus providing fresh evidence on long term trends of income mobility for individuals belonging to different cohorts in the same phase of their life; (iii) we exploit mobility and volatility estimates to study the relationship between mobility as a linear trend and volatility as earnings fluctuations, even when moderated by the direction of the trend.

To the best of our knowledge, few analyses – and none for Italy – investigated the components of income dynamics observing individuals for a long period (e.g., Nichols and Rehm 2014, Latner 2018) and no studies have carried out this analysis comparing a large number of workers cohorts to provide evidence on the long-run trend of income dynamics in a 50-year period. We thus apply a cohort approach to study what happens within groups that are quite homogeneous with respect to structural conditions, and to compare these groups with each other to understand how the phenomena under study have evolved over time.

In more detail, applying various methods (see Section 5), we descriptively characterize patterns of inequality, mobility and volatility in Italy across several cohorts of workers, to understand whether the well-proved increase in labour market inequality experienced by Italy has been compensated by higher mobility between workers or has been due to widening persistent differences. Moreover, we go into details of income changes to distinguish 'good' mobility – i.e., upward and predictable changes – from mere volatility – i.e. frequent and unpredictable fluctuations.

Even if we do not make social welfare evaluations in this work, we start from the idea, discussed in Jenkins (2011), that income mobility is 'good' for society if there is less association between starting and ending income, because it means being closer to equality of opportunity as a concept of 'origin independence'. This is fairly accepted as true in the intergenerational context, more complex in the intragenerational one: how to establish whether moving away (upward or downward, more or less frequently) from one's own starting point is good or not? Surely one can consider intragenerational mobility desirable in its role of reducing permanent income inequality. If there is mobility, the income of one individual for any given year cannot coincide with his income averaged over several years, so mobility can be seen as a measure of longitudinal variability. However, mobility from one's own starting position is 'bad' when it means frequent income fluctuations. As argued in OECD (2018), 'unequal mobility' can occur when unpredictable income changes combine with low levels of long-term (upward) income mobility and when this concerns mostly the most vulnerable population groups.

Therefore, a rich and comprehensive picture of the evolution of earnings dynamics must bring together elements of inequality, mobility, and volatility. In Section 2 we provide a conceptualization of these three notions, highlighting their point of intersection with the literature on income dynamics. Then, using the rich information contained in our dataset, we apply to each cohort different measures of mobility as income growth and positional change, and then decompose income changes into movements that are smooth and directional and unpredictable fluctuations (Nichols and Rehm, 2014; Latner, 2018). This decomposition allows to evaluate the relative contributions to total inequality of permanent differences among individuals, individual mobility and volatility, and to study in a regression framework the relationship between mobility and volatility. Moreover, repeating the decomposition distinguishing workers according to their main features, we assess differences in income patterns according to worker education and area of work and also comparing in some analyses males and females.

In detail, the paper is structured as follows. Section 2 briefly reviews the main concepts used in the economic literature to assess income dynamics and motivates the choice of the concepts we follow in this article. Section 3 provides a general description of the structural changes in the Italian labour market from the 1970s. Section 4 presents the characteristics of our dataset, while Section 5 discusses the various empirical methods we use to investigate the phenomena under scrutiny. Section 6 shows our main results and Section 7 concludes.

2. Related literature *(very preliminary and incomplete draft)*

Various concepts have been used in the economic literature to deal with issues related to individual income changes over time. We first treat separately the concepts of income mobility, intragenerational inequality, and earnings insecurity, and then explore how they can intersect.

Income mobility can be defined in terms of distance between origin and destination income, where the destination must be defined at a later time than the origin. The change can be defined either with reference to one's own starting point (structural mobility), in absolute or relative terms, or as change in the position in the income distribution relative to others (positional mobility). There are three dimensions to consider: the cross-sectional origin situation, the cross-sectional destination situation, and the longitudinal dimension that links the two.

As discussed in Section 5, in the core analyses of this article we rely on the descriptive measure of income trajectories – developed according to the 'income trend framework' – firstly proposed by Nichols (2008) to decompose income dynamics into long-run inequality, mobility (i.e. dispersion in individual-specific growth rates) and volatility (i.e. dispersion in intertemporal deviation around individual-specific growth rates). Following Latner (2018), we then use findings from this decomposition to estimate the link between individual mobility and volatility.

This decomposition method has been firstly applied by Nichols (2008) to investigate, through PSID data, trends of household income in the US from 1970 to 2005 following individuals aged 30-60 longitudinally for a 9-year period. Tracking households for 9 calendar years, Nichols (2010) has then compared household income mobility trends between the US and China from 1989 to 2005. Nichols and Rehm (2014) analysed trends in household income inequality components in Canada, Germany, Great Britain, and the US from 1986 to 2009, also providing some results for 26 additional countries for a shorter 4-year time span. Finally, Latner (2018) used the PSID survey biannual data from 1970 to 2013 observing household incomes for an 11-year period with the main aim of investigating the link between mobility and volatility.

Decompositions based on the income trend framework – despite their advantage of providing statistics for the three components of income dynamics – have then been scarcely applied in the empirical literature. As explained in Section 5, we rely on this framework, advancing with respect to the literature from various perspectives: (i) differently from the quoted studies, we focus on individual earnings rather than on household incomes; (ii) instead of comparing over time trends of individuals belonging to a certain (large) age class, we follow a cohort-style approach and compare at the same phase of their working life the income dynamics patterns of individuals born in 7 subsequent 5-year birth cohorts; (iii) we focus on Italy, an EU country characterised by an intense process of labour market and wage-setting reforms in the last decades.

3. Background: structural changes in the Italian labour market from the 1970s *(to be added)*

4. Data

We use the AD-SILC panel dataset, developed merging the 2004–2017 waves of the IT-SILC survey (the Italian component of the EU-SILC) with the administrative longitudinal social security records collected by the Italian National Social Security Institute (INPS). The INPS archives record employment and earnings histories of all individuals working in Italy from the moment they enter the labour market – even if reliable earnings data are available from 1974 for the employees in the private sector – up to 2018. In addition to the demographic characteristics, the administrative component allows us to have detailed information on the type of employment contract, the gross annual earnings, the weeks worked in the year.

This dataset is particularly suited for our analysis, because of two characteristics that are crucial and rare in the existing literature on income dynamics: (i) workers are followed every year for a large part of their career, allowing us to distinguish between short and long-term dynamics; (ii) they are followed continuously as long as they participate in the formal labour market – without memory biases and, mostly, the 'gaps' from attrition characterizing panel data from surveys. This last characteristic largely improves the analysis: volatility is traditionally considered to be a short-term issue and requires observations very close in time, while mobility can be studied both as a short and a long-term phenomenon. Having a long span of income records without "holes" allows us to study mobility and volatility at the same time looking also at their interaction.

The sample is restricted excluding individuals without Italian citizenship, since the retrospective panel under-represents them in older cohorts. We focus on men in our baseline analysis to get rid of issues related to gender differences in labour market participation. However, we perform all the analyses also on the women sample in the appendices. We also focus on those working as employees in the private sector, even if we use information about periods spent working in the public sector or as a self-employed to impute unemployment spells[1].

We follow individuals for the 11-year period when they are aged 35-45[2]. We run our analysis both on the subsample of individuals with positive earnings from private employment in the whole 11-year period and on the full sample where also individuals with zero earnings in few years are considered in the income dynamics. However, being interested at assessing mobility and volatility, we exclude from the analyses also those individuals characterized by very long periods as 'zero earners', namely those (0.49%) with total unemployment spells in the 11-year period longer than 4 years.

We use as outcome variable real annual earnings including allowances (price level 2015), gross of personal income taxes and social contributions paid by the employee. We focus on annual

[1] For the imputation of unemployment periods, we exploit several sources of information: thanks to the INPS archive we know whether the worker is receiving unemployment benefits and whether or not he or she is part of the formal labor market, whatever the job position. In addition, thanks to the EU-SILC component we know at what age the worker finished studying (assuming that he or she is available for work from that moment).

[2] The very few individuals out of the private dependent market only when 35-years-old are followed when aged 36-46. They are 0.16% of the baseline sample, and 0.14% of the sample which includes zero earners. For females, the two percentages are 0.29% and 0.26%.

earnings rather than on monthly or weekly wages,[3] since we are mostly interested in individual monetary well-being, which depends on total income received by an individual in a certain year as a combination of hourly wage and weeks worked in the year.

As mentioned, for the baseline analysis we restrict the sample to male workers only, for two main reasons: (i) the focus of the analysis is the description of patterns of income inequality, mobility, and volatility across cohorts, and we have relevant reasons to believe that men and women differ in the mechanism underlying such income phenomena; (ii) the abovementioned issue is even more relevant when including imputed unemployment periods in the analysis. The baseline analysis for male workers does not include periods with zero earnings, which are included later on for comparison.

The analysis is performed on workers aged from 35 to 45 years, divided into seven five-year cohorts of birth: 1940-1944, 1945-1949, 1950-1954, 1955-1959, 1960-1964, 1965-1969, 1970-1974. Thus, we observe earnings patterns from 1975 to 2018.

The AD-SILC dataset has an age composition issue due to its construction method: since the first EU-SILC interview took place in 2004, we can observe only very few workers who were in the last phase of their career in the first years of the period covered by our INPS data (from 1974 on). Therefore, the share of older workers in the dataset increases during the whole period. To address this issue, we restrict our analysis to the 35-45 age window, so that (i) the number of workers changes over time according to population and labour market dynamics, rather than dataset construction mechanics; (ii) we can compare workers observed during a common 11-year life-cycle stage.

As highlighted in Haider and Solon (2006), comparing the current income of workers at different stages of their career may introduce a non-negligible bias in every kind of analysis involving earnings measures. This consideration enforces our choice of selecting a common stage of life for every cohort. We also know that the selected window can approximate the lifetime earnings experience of the workers included in the analysis. The literature on intergenerational inequality (Haider and Solon 2006, Böhlmark and Lindquist 2006), respectively for the US and Sweden, finds evidence that the difference between yearly and lifetime earnings for men is minimized when individuals are observed in their median age, i.e. around age 35[4]. Conversely, a simple rule does not emerge for women, who display more variety in their life-cycle income patterns. We provide a simple descriptive verification of the suitability of our selected age window in Appendix A, concluding that in our sample, on average, current earnings coincide with lifetime earnings at some point between ages 35 to 40, regardless of gender and education. Finally on this point, Nybom and Stuhler (2016) warn that age-earnings profiles may be worker-, country- or cohort-specific even for male workers, so that the choice of the same point in age for every worker may be misleading. For this reason, our choice of looking at workers' outcomes when they are between the ages of 35 and 45 can help in studying their lifetime possibilities without making too

[3] INPS data do not record hourly wages.

[4] The issue of finding the best moment in life to approximate lifetime income is particularly relevant for the intergenerational inequality literature. The reason is that, when analysing the effect of parents' characteristics on children outcome, it is crucial not to disregard in which stage of life children are observed and their outcome evaluated.

strong assumptions and generalizations. Moreover, an age window of 11 years is sufficiently short to allow the use of linear models to identify individual income trends without the need for log transformations (we also use a quadratic trend as a robustness check).

As a standard procedure with administrative data, the bottom and top 0.5% of the earnings distribution in each year by gender are dropped to minimize measurement errors that may occur at the tails. In the literature on income mobility and earnings volatility, trimming the distribution at the top is often necessary also to exclude serious outliers to which mobility and volatility measures are very sensitive.

Since for the income risk decomposition we need a balanced sample to avoid bias in the estimation of volatility, we present below descriptive statistics for the sample including male workers that are observed at every age from 35 to 45. When balancing the panel of men (women), we keep 68.43% (51.15%) of workers, 70.37% (52.22%) for the sample including zeros. Summary statistics for women are reported in Appendix B.

Table 1: Summary statistics of real annual earnings (€) by cohort of birth

Cohort of birth	Period	N	Mean	SD	p10	p50	p90	Gini	GE2
1940-1944	1975-1989	25322	25767	9543	16600	23851	38179	0.195	0.069
1945-1949	1980-1994	29832	28487	10714	18178	26524	41605	0.196	0.071
1950-1954	1985-1999	26818	30027	11560	19169	27427	45173	0.201	0.074
1955-1959	1990-2004	28787	30367	12406	18640	27565	46567	0.214	0.083
1960-1964	1995-2009	33583	29982	13677	17295	26648	47956	0.238	0.104
1965-1969	2000-2014	38302	30483	14334	17096	27293	48474	0.242	0.111
1970-1974	2005-2018	28039	31769	15009	17474	28511	50804	0.244	0.112

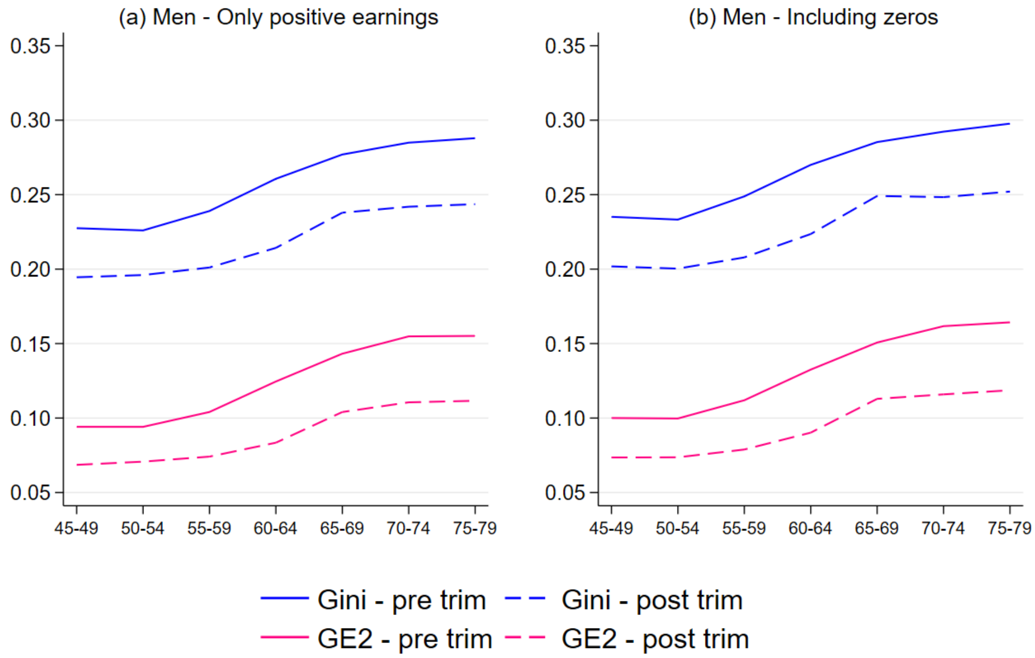
Table 2: Summary statistics of real annual earnings (€) by cohort of birth, including zeros

Cohort of birth	Period	N	Mean	SD	p10	p50	p90	Gini	GE2
1940-1944	1975-1989	26037	25448	9759	16060	23675	37976	0.202	0.074
1945-1949	1980-1994	30382	28268	10848	17883	26397	41477	0.200	0.074
1950-1954	1985-1999	27731	29614	11760	18687	27186	44803	0.208	0.079
1955-1959	1990-2004	29755	29838	12673	17813	27241	46202	0.224	0.090
1960-1964	1995-2009	34848	29351	13947	15993	26298	47554	0.249	0.113
1965-1969	2000-2014	39567	30049	14466	16288	27021	48141	0.248	0.116
1970-1974	2005-2018	29447	31134	15167	16282	28089	50180	0.252	0.119

The final male sample is composed of 19,153 (19,797) workers followed for 11 consecutive years in their mid-career excluding (including) zero earnings.

To evaluate the impact of reducing the sample to workers followed continuously, we compare the Gini and the GE2 index before and after the balancing procedure. Each index is computed within-cohort, and the graphs are reported in Figure B for men, and in Appendix B for women.

Figure 1: Impact of panel balancing, men only



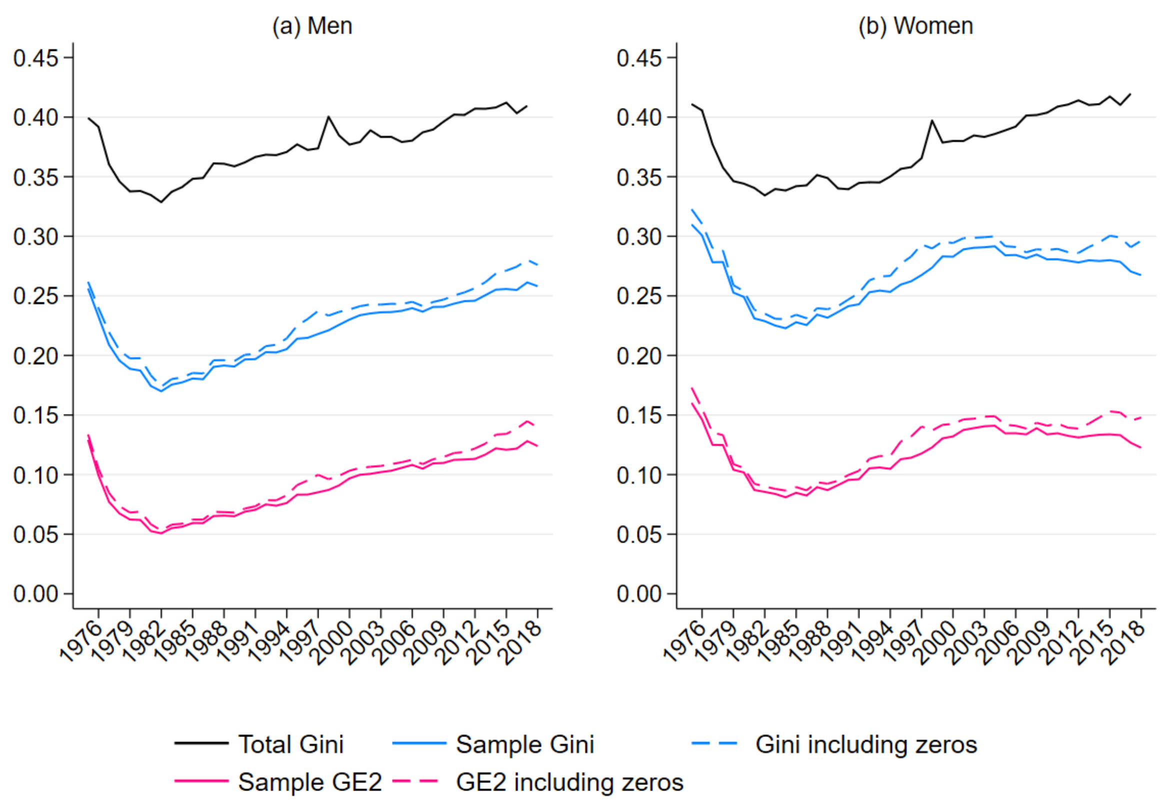
The solid lines plot the inequality indices by cohort for workers aged from 35 to 46. The dashed lines plot the inequality indices by cohort computed excluding workers who are not present for 11 consecutive years in the sample.

We see from the figures that the pattern of the two indices is not dramatically influenced by the reduction, for both men and women, both including or removing the zeros. As expected, within-cohort inequality in increasing from the second cohort on and higher when including zero earners. When we balance the panel, we lose (i) workers who come from or move to public of self-employment when they are aged from 35 to 45, and (ii) workers who have more than four periods of unemployment in that age window. Moreover, in the case of panel (a) of Figure 1, we also lose all workers experiencing at least one period of unemployment. This means that we are selecting 'stronger' workers, those who are more attached to the labour market. This is another important reason to keep separate the analysis for men and women, given their different (and differently evolving) participation and attachment. As we see in Figure 1, selecting stronger workers we reduce the level of overall inequality but the pattern is similar: the solid and dash lines slightly diverge over time, suggesting that the number of 'weak' workers excluded is higher for more recent cohorts. Finally, it is interesting to notice that the choice of the GE2 index to measure inequality affects the level but not the pattern of inequality with respect to the more traditional Gini index.

We evaluate the suitability of our sample for the analysis by looking at the trend of the two inequality indices over time: Figure 2, based on administrative data on Italian private employees, reports a comparison of the Gini and the GE2 indices computed by gender for the sample 35-45 and the overall Gini measured using the entire 15-64 population of workers in the Italian private dependent sector, again by gender.

It can be noticed that (i) the patterns of the Gini and the GE2 indices for the 35-44 sample are very similar within-gender, while the levels are different; (ii) the pattern of inequality over time when considering only 35-45 workers is similar to the general one for men, while it is diverging from mid-2000s for women. These patterns suggest that the impact of selecting only workers with no more than four zeros has a strong impact on the female sample excluding those with lower earnings, with the result of lowering inequality in the period in which we expect it to rise after the Great Recesssion.

Figure 2: Inequality patterns in the Italian private dependent sector from 1975 to 2018



Total Gini is computed for the entire 15-65 population of employees in the private sector using INPS data

The use of imputed zeros in part of the analysis allows us to mitigate the potential underestimation of the part of volatility that is due to unemployment periods. It can be seen in Figure 2 that, for both Gini and GE2 indices, the pattern of inequality without zeros diverges from that including zeros in periods of recession, with two clear ‘diverging’ points in 1992 and 2008.

5. Methods

Our description of the patterns of earnings inequality, mobility, and volatility across cohorts is based on comparisons between ‘peers’: every worker is observed when he is between 35 and 45 years-old, independently of the calendar period. We compare within-cohort income inequality, mobility, and volatility using different definitions of income change to have a broad picture of how workers have moved and how such movements have changed over time.

In our investigation, we follow three main steps. First, we provide some synthetic measures of income mobility of each cohort in the 11-year period in terms of absolute and positional mobility. Second, we follow Nichols (2008) and Nichols and Rehm (2014) applying an inequality decomposition in three components (permanent inequality, mobility risk, and volatility), according to the 'income trend framework'. Finally, using the individual measures of mobility and volatility provided in the second step, we explore the relationship between volatility and mobility in a regression framework.

As mentioned, we carry out our decomposition exercise considering both only those employees with positive earnings in the age window considered and also adding those who experience some years with 'zero earnings'. Moreover, we distinguish workers by education and geographical area of work, allowing for the possibility of different income mobility patterns across subgroups.

5.1 Synthetic measures of income mobility

To describe the patterns of income mobility in terms of structural change and positional mobility, we compare the origin and destination income of each worker with those of the other workers who belong to the same cohort and are in the same stage of life. In our first empirical approach, the 'origin income' is defined as earnings averaged from age 35 to 37, while the 'destination income' as earnings averaged from age 43 to 45.[5] Averaging income in a short interval to slightly smooth it is the standard procedure to build mobility measures mitigating the effect of year or age-specific shocks. Given the above definitions of origin and destination income, we compute some measures of absolute and relative mobility for each cohort. Moreover, to account for the differences that surely characterize the patterns of mobility at different parts of the earnings distribution, we report mobility measures by the starting income decile. Note that the income decile is computed with reference to the starting distribution, fixing ages rather than time: the position of each worker is relative to that of the other workers who belong to the same cohort and are aged from 35 to 37, independently of the calendar year.

We select three synthetic measures that can help in understanding mobility patterns: (i) the average earnings growth as a percentage of starting income; (ii) upward and downward positional mobility, captured by the share of workers moving to a higher or a lower earnings decile from the origin to the destination distribution; (iii) the average number of 'jumps' – deciles crossed – when moving up or down.

5. 2 Income risk decomposition

To distinguish permanent inequality from mobility, and 'good' mobility from volatility, we follow the idea of Nichols (2008) and Nichols and Rehm (2014) of measuring overall inequality through a subgroup decomposable index, using individuals themselves as the population subgroups. The subgroup decomposability of an inequality index (Shorrocks, 1980, 1984) requires that the within-group observations share common characteristics, and that the overall index does not depend on the selection of the subcategories. The Generalized Entropy Index with parameter 2

[5] For those few individuals which are followed from age 36 to 46, the origin income is defined as income averaged from age 36 to 38, while the destination income as income averaged from age 44 to 46.

(GE2; half the squared coefficient of variation) shares these characteristics with the family of Generalized Entropy indices, and requires considering earnings without log transformation allowing the inclusion of zeros:

$$GE(\alpha) = \frac{1}{N}\,\frac{1}{\alpha(\alpha-1)}\left[\sum_{i=1}^{N}\left(\frac{x_i}{\mu}\right)^{\alpha} - 1\right],\ \alpha \neq 0{,}1, i = 1, \dots, N \tag{1}$$

$$GE(2) = \frac{1}{2}\left(\frac{\sigma}{\mu}\right)^2 \tag{2}$$

Applying a decomposition by subgroups to panel data for a panel of $i = 1, \dots, L$ workers followed for $t = 1, \dots, T$ periods, for a total of N=LT observations, the 'between-group' inequality component measures permanent inequality across workers – i.e. inequality in average incomes over the observed time window -, while the 'within-group' inequality component measures average 'personal inequality' over time, a combination of mobility risk and volatility. Let y_{it} be annual real gross earnings of worker i at age t, $\bar{y}$ the average annual earnings among all N=LT observations in the age window T for the L workers of the cohort, and $\bar{y}_i$ average earnings of worker i in the T periods. Separately for each cohort, we decompose overall within-cohort inequality as follows:

$$GE(2) = \frac{1}{2\bar{y}^2}\left[L^{-1}\sum_{i=1}^{L} T^{-1}\sum_{t=1}^{T}(y_{it} - \bar{y})^2\right] \tag{3}$$

$$= \underbrace{\frac{1}{2\bar{y}^2}\left[L^{-1}\sum_{i=1}^{L} T^{-1}\sum_{t=1}^{T}(\bar{y}_i - \bar{y})^2\right]}_{\text{“Between-workers” inequality (B)}} + \underbrace{\frac{1}{2\bar{y}^2}\left[L^{-1}\sum_{i=1}^{L} T^{-1}\sum_{t=1}^{T}(y_{it} - \bar{y}_i)^2\right]}_{\text{“Within-worker” inequality (W)}}$$

The 'between workers' inequality is the variance of individual-level mean income, divided by twice squared mean income. It corresponds to the definition of long-term inequality as the dispersion in individual income averaged along the time window considered. The 'within-worker' inequality is the mean across workers of the individual-level variance of income over time, again divided by twice squared mean income. We do not need to weight the 'within' variances since all the individuals are observed for the same number of years. The presence of average earnings squared at the denominator makes it easier to compare the results for different cohorts because it removes the effect of overall income level from every measure. As a further and crucial step, the numerator of the within-worker inequality component can be further decomposed into 'mobility risk' (average dispersion of worker-specific trends) and volatility (average individual variability around a personal trend).

We can make the decomposition relying on a regression of the type

$$y_{it} = \alpha_{0i} + \alpha_{1i}t + e_{it} \tag{4}$$

of current income y_{it}on an age trend t with age centered at zero so that α_{0i} is equal to individual mean income in the window, and no assumption on the distribution on the error term so that the method is fully descriptive. Since $\alpha_{0i} = \overline{y_i}$ by construction, substituting Equation (4) in the within-worker component of inequality, we obtain:

$$W = \frac{1}{2\bar{y}^2}\left[L^{-1}\sum_{i=1}^{L}T^{-1}\sum_{t=1}^{T}(y_{it}-\overline{y_i})^2\right] = \frac{1}{2\bar{y}^2}\left[L^{-1}\sum_{i=1}^{L}T^{-1}\sum_{t=1}^{T}(\alpha_{0i}+\alpha_{1i}t+e_{it}-\overline{y_i})^2\right]$$

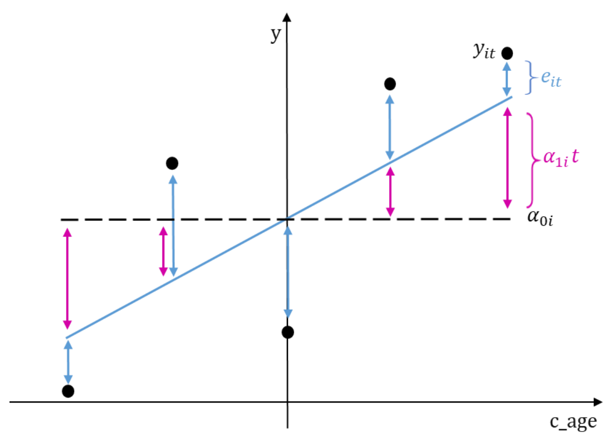
$$= \frac{1}{2\bar{y}^2}\left[L^{-1}\sum_{i=1}^{L}T^{-1}\sum_{t=1}^{T}(\alpha_{1i}t+e_{it})^2\right] = \frac{1}{2\bar{y}^2}\left[L^{-1}\sum_{i=1}^{L}T^{-1}\sum_{t=1}^{T}((\alpha_{1i}t)^2 + {e_{it}}^2 + 2(\alpha_{1i}t)(e_{it}))\right]$$

$$= \frac{1}{2\bar{y}^2}\left[L^{-1}\sum_{i=1}^{L}T^{-1}\sum_{t=1}^{T}(\alpha_{1i}t)^2 + L^{-1}\sum_{i=1}^{L}T^{-1}\sum_{t=1}^{T}{e_{it}}^2 + L^{-1}\sum_{i=1}^{L}T^{-1}\sum_{t=1}^{T}2(\alpha_{1i}t)(e_{it})\right]$$

$$= \underbrace{\frac{1}{2\bar{y}^2}\left[L^{-1}\sum_{i=1}^{L}T^{-1}\sum_{t=1}^{T}(\alpha_{1i}t)^2\right]}_{\text{Mobility risk}} + \underbrace{\frac{1}{2\bar{y}^2}\left[L^{-1}\sum_{i=1}^{L}T^{-1}\sum_{t=1}^{T}{e_{it}}^2\right]}_{\text{Volatility}} + \underbrace{\frac{1}{2\bar{y}^2}\left[L^{-1}\sum_{i=1}^{L}T^{-1}\sum_{t=1}^{T}2(\alpha_{1i}t)(e_{it})\right]}_{\text{Residual component}}$$

Figure 3 below shows an example of the different components of income.

Figure 3: Earnings components in an 'income trend' framework



The choice of a linear trend, which may be controversial when considering the entire life-cycle income pattern, is considered particularly suitable when looking at incomes in a short age window such as 35-45 years old.[6]

We show results based on a time window of 11 years for each worker, but we also carried out the analyses considering narrower time windows. As discussed in Nichols (2008), the length of the period must be at least three (two observations to estimate a linear trend, and the third to allow deviation from it). However, the variance of the idiosyncratic error term used to characterize volatility will tend to be dramatically understated for small lengths. Nichols (2008) uses T = 5 and Nichols and Rehm (2014) use T=3, while we decide to adopt a wider time range T=11. Our results are robust to the use of shorter windows (5, 7 and 9 years; results are reported in Appendix D).

[6] In robustness checks available upon request we also performed analyses assuming a quadratic trend and our main results do not change.

As a further caveat, part of residuals' behaviour may be due to measurement error; however, since we estimate inequality components for different birth cohorts, we consider reliable our analysis as long as measurement error has not changed over time.

We build confidence intervals for the decomposition estimates using standard errors from bootstrap sampling (1000 iterations) and standard normal based thresholds.

5.3 Relationship between mobility and volatility

The final part of our empirical analysis is dedicated to exploring the link between income volatility and mobility across cohorts, as measured using the income risk decomposition discussed in Section 5.2. Distinguishing changes that are 'smooth and directional' (Latner 2018) from those that are volatile allows to analyse the relationship between these two kinds of changes, trying to understand, on the one hand, whether steeper trends are associated to more volatile incomes and, on the other hand, whether the link between mobility and volatility differs according to the direction of the trend.

To analyse this relationship we apply an OLS linear regression of earnings volatility on earnings mobility at the individual level. Each worker has a single measure of volatility defined as the mean squared residual from a personal linear trend, and a single measure of mobility defined as the mean squared trend in the 11-years window. Thus, we have no longer a panel structure and make use of an OLS model of the type:

$$Volatility_i = \beta_0 + \beta_1 Mobility_i + \beta_2 Upward_i + \beta_3 Origin\ income_i + \beta_4 Origin\ decile_i + error_i \quad (5)$$

fitted separately for each cohort of birth to allow the relationship to change over time. Volatility and mobility are expressed as the natural logarithm of the original individual measures, so as to approximate a normal distribution and to be able to interpret the coefficient β_1 as an elasticity.

To allow volatility to be differently related to mobility depending on the direction of the trend, we also introduce an interaction term between the mobility estimate and the dummy for the upward trend:

$$Vol_i = \beta_0 + \beta_1 Mobility_i + \beta_2 Upward_i + \beta_3 Origin\ income_i + \beta_4 Origin\ decile_i + \\ + \beta_5 Mobility_i * Upward_i + error_i \quad (6)$$

Moreover, in a further specification we control for worker's characteristics, i.e. education, area of work, average number of weeks worked in a year and broad occupational qualification (blue-collar, white-collar and manager).[7]

[7] Detailed occupational groups are not recorded in administrative archives, and we cannot take this information from EU-SILC since it only reports the occupation at the time of the interview.

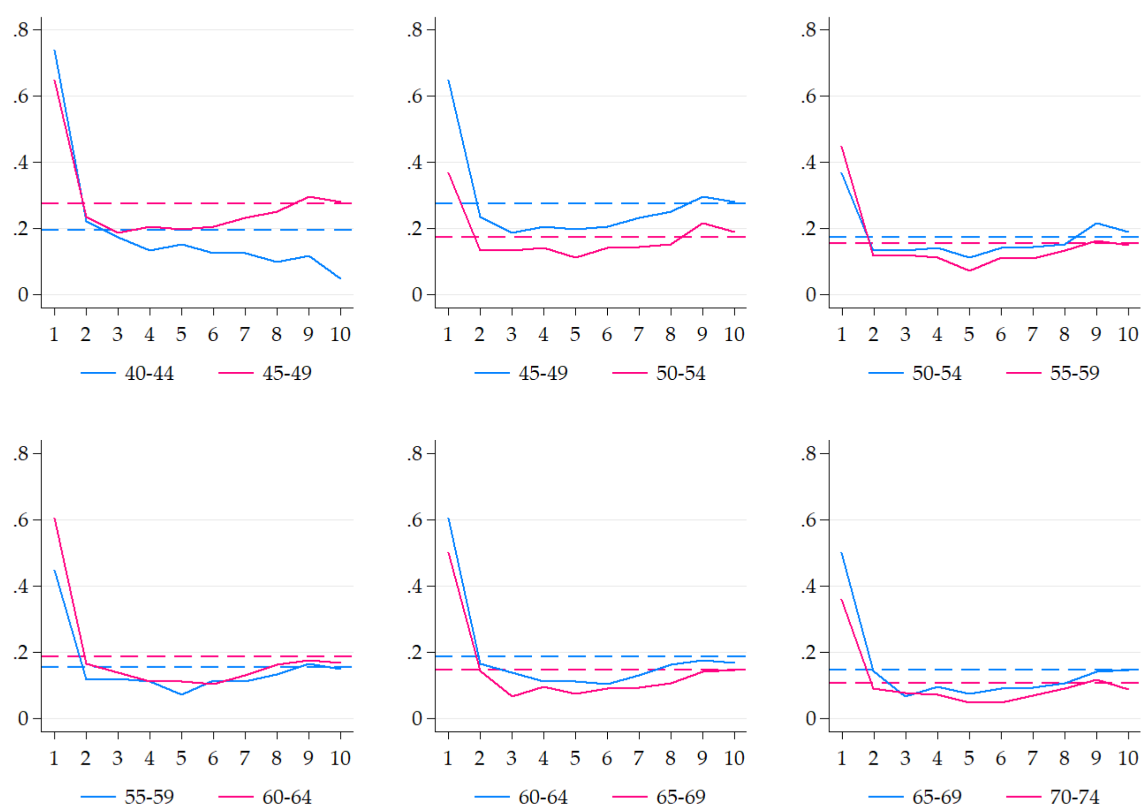
6. Results

6.1 Synthetic measures of income mobility

We present in this section the results for the baseline sample made of men excluding zero earners. The results for female workers are in Appendix B, while those for men including zero earnings and divided by education and area of work are in Appendix C.

We show in Figure 4 the average income growth in percentage terms by within-cohort starting decile. The seven cohorts of birth are compared in consecutive pairs to appreciate changes from one cohort (in blue) to the next (in red). The dashed lines report the average growth irrespective of the starting decile. As expected, the lowest decile experiences the higher growth in relative terms; in a mean-reverting 'pro-poor' growth context (the higher the starting income, the lower the growth rate), we should expect a decreasing curve from the bottom to the top decile. On the contrary, this happens only for the first cohort; since it is observed from 1975 to 1989 (when the workers born from 1940 to 1944 are between ages 35 and 45), it is the only one to be completely covered by the 'Scala mobile' system of wage indexation against inflation. This system was based on allowances equal for everybody in absolute terms, thus providing higher real-wage protection for workers at the bottom of the wage distribution. It was reformed towards proportionality through the 80s until complete abolition in 1993.

Figure 4: Relative income growth (%) by within-cohort starting decile



Beyond the first cohort, for all other cohorts we see an L-shape deformed on the right, with the highest deciles experiencing greater percentage growth than the middle deciles, despite starting from higher income.

Looking at the dynamics of positional mobility, Figure 5 shows that, while the average distribution of upward mobility by deciles mainly reflects the possibilities to go up (the lower the position, the higher the opportunity to rise), that of downward mobility goes in the opposite direction: from the fifth decile on, there is a sheltering mechanism that prevents those in higher positions from descending. The higher the starting point, the less risk of ending up in a lower position after 11 years during the mid-career.

Figure 5: Fraction of workers changing position, by within-cohort starting decile

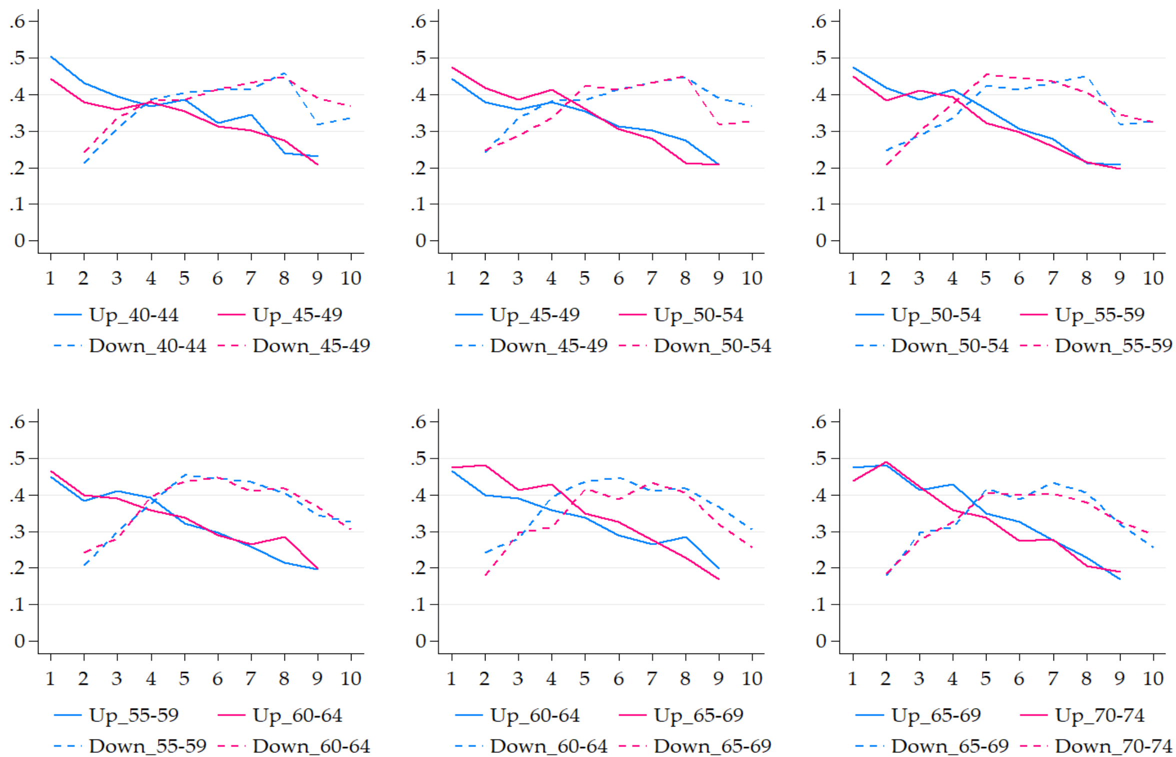
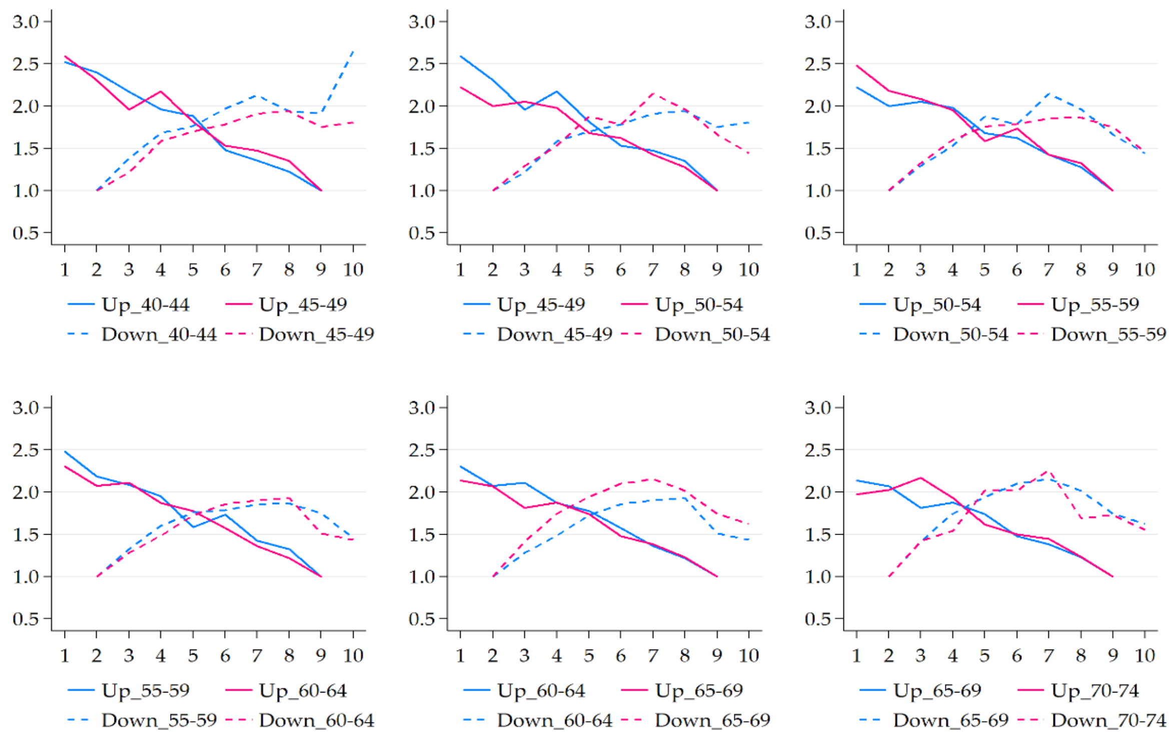


Figure 6: Average 'jump' of workers that change position, by within-cohort starting decile



Finally, as concerns positional mobility, we can also look at how far workers move on average when changing position: comparing upward and downward movements in Figure 6, we see that, with the only exception of the first cohort, there is an asymmetry in the distribution of the distance depending on the direction. Again, high-positioned workers are, on average, more protected from moving away from their starting position.

6.2 Income risk decomposition

The detailed results of the income risk decomposition are reported in Appendix C, while Figure 7 and 8 summarise the main results by birth cohort. The general rise in total within-cohort earnings inequality in the baseline sample is confirmed and it is coupled with a rising contribution of permanent, between-workers inequality. Moreover, the 'within-worker' component is almost equally divided into a mobility component and a volatility component from the third cohort onwards except for the second cohort that enjoys 'good' mobility more than the others.

Figure 7: Income risk decomposition, by cohort of birth

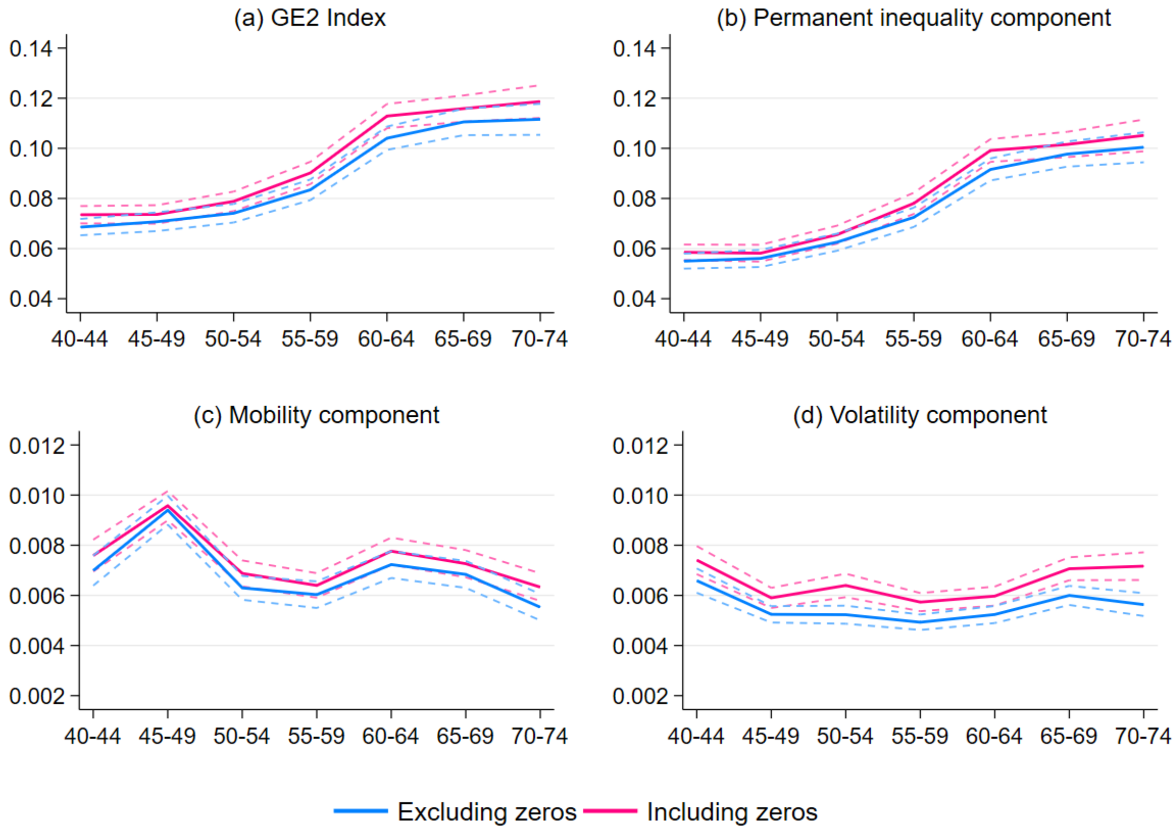
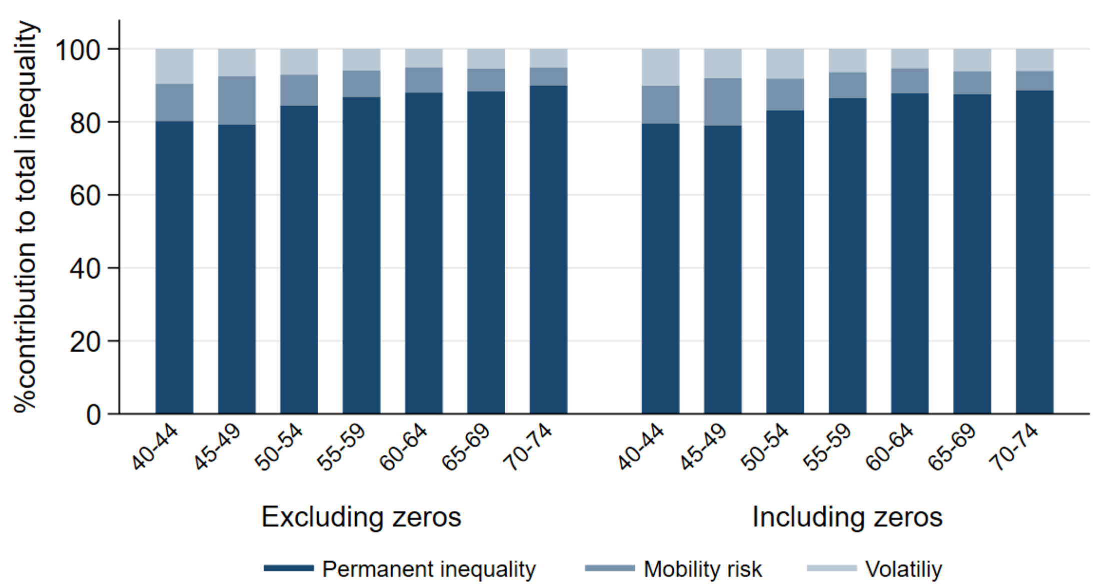


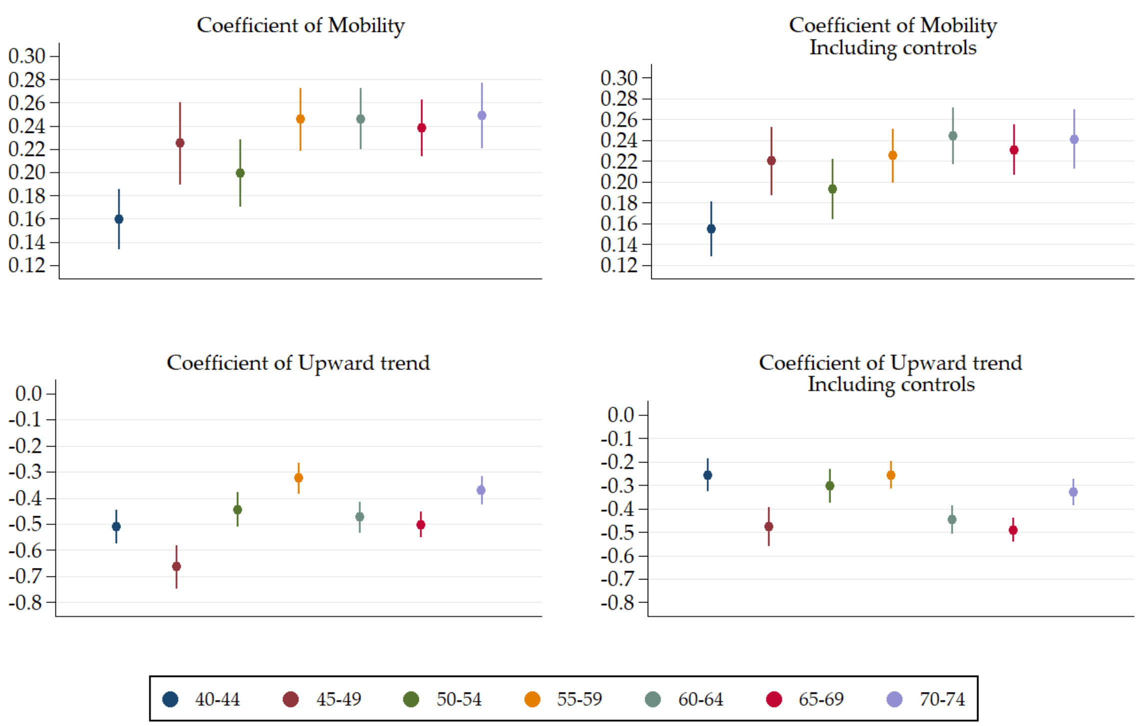
Figure 8: Contribution to overall inequality, by cohort of birth



6.3 Relationship between mobility and volatility

When studying the relationship between earnings mobility and volatility in mid-career, we firstly notice (Figure 9) that a 1% increase of the steepness of the linear trend is associated with a positive but lower than one (between 0.16 and 0.25, depending on the cohort) increase in volatility. Steeper trends, with the same income and starting position, as well as trend direction, are associated with higher volatility. Looking at the coefficients of the dummy for upward trend, we see that they are all significant and negative: all other things being equal, including the slope of the trend, upward trends are associated with lower volatility.

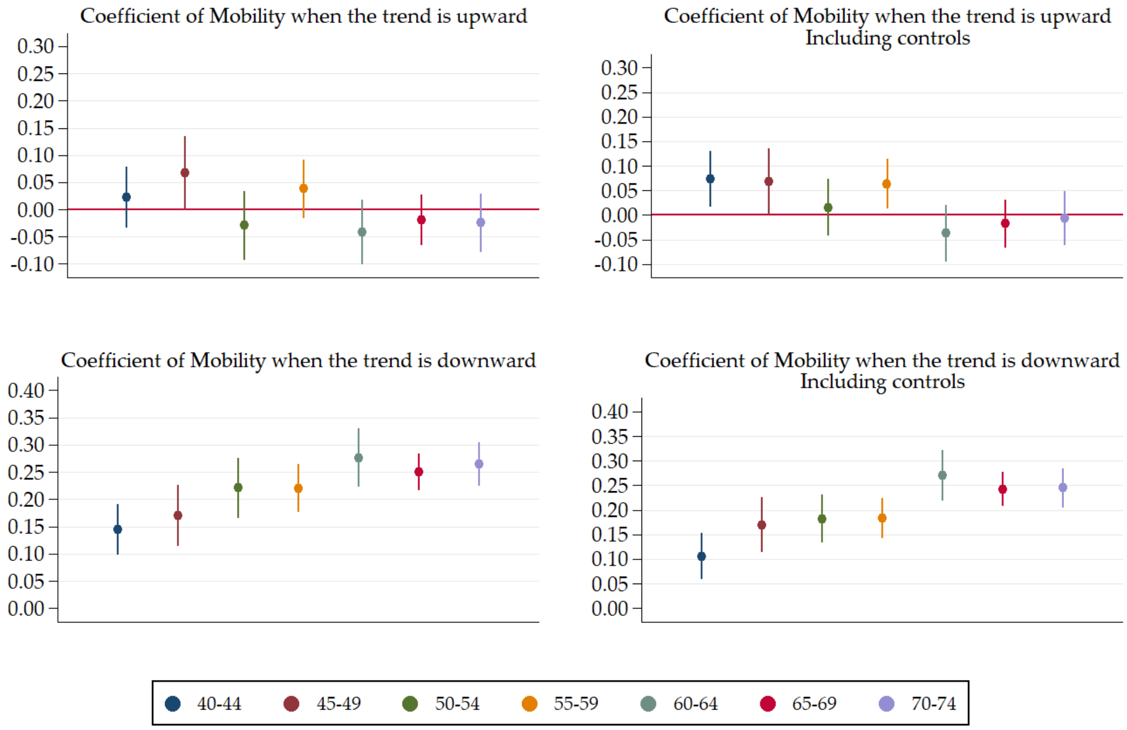
Figure 9: Elasticity of volatility with respect to mobility, by cohort of birth



The baseline model regresses volatility (log of the standard deviation of individual residuals from a linear trend) on mobility (log standard deviation of individual trend points), starting earnings, starting decile, and a dummy for upward trends. The controls include highest level of education, 3-level qualification, prevalent area of work, average number of weeks worked in a year

When including in the model an interaction between the dummy for upward trend and the mobility measure, we see from Figure 10 that the relationship between mobility and volatility is not statistically significant when the trend is upward sloped, while it becomes significant and positive (between 0.15 and 0.27) when the direction is downward. Moreover, as concerns the latter coefficient, we see a rising trend across cohorts from the first to the fifth, and a flattening thereafter.

Figure 10: Elasticity of volatility with respect to mobility including an interaction term, by cohort of birth



The baseline model regresses volatility on mobility, a dummy for upward trends, an interaction between the two, starting earnings, starting decile. The controls include highest level of education, 3-level qualification, prevalent area of work, average number of weeks worked in a year

Having a positive trend has an overall negative effect on volatility, all else being equal; moreover, if we look at the effect of the steepness of the trend on volatility, having a positive trend makes the relationship not statistically different from zero, while when the trend is negative the relationship is positive and significant. These results, which are robust to the inclusion of various controls, suggest that volatility is a threat especially for workers who exhibit a negative trend, who are already more vulnerable than the others. Moreover, the results do not change notably using the sample with zeros included.

7. Conclusions *(to be added)*

References

Atkinson A.B. (2015), *Inequality. What can be done?*, Cambridge University Press.

Böhlmark, A. and Lindquist, M.J. (2006), *Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden,* Journal of Labor Economics, Volume 24, Number 4.

Burkhauser R., Couch K. (2009), *Intragenerational inequality and intertemporal mobility*, in Salverda W., Nolan B., Smeeding T. (eds), *Handbook of economic inequality*, Oxford University Press.

D'Ambrosio, C. (2018), *Handbook of Research on Economic and Social Well-Being,* Edward Elgar Publishing.

Franzini M., Raitano M. (2019), *Earnings inequality and workers' skills in Italy*, *Structural Change and Economic Dynamics*, vol. 51, pp. 215-224.

Haider S., Solon G. (2006), *Life-cycle variation in the association between current and lifetime earnings*, American Economic Review, 96 (4), pp. 1308-1320.

Jenkins, S. (2011), *Changing Fortunes: Income Mobility and Poverty Dynamics in Britain*, Oxford University Press.

Latner, J.P. (2018), *Income volatility and mobility: A conceptual exploration of two frameworks*, Research in Social Stratification and Mobility, Volume 53, Pages 50-63.

Nichols, Austin, (2008). *Trends in Income Inequality, Volatility, and Mobility Risk*, IRISS Working Paper Series 2008-10, IRISS at CEPS/INSTEAD.

Nichols, Austin, (2010). *Income inequality, volatility, and mobility risk in China and the US*, CES 2009, China Economic Review, Volume 21, Supplement 1, pages S3-S11.

Nichols, Austin, and Philipp Rehm, (2014). *Income Risk in 30 Countries,* Review of Income and Wealth, International Association for Research in Income and Wealth, vol. 60(S1), pages 98-116, May.

Nybom, M. and Stuhler, J. (2016), *Heterogeneous Income Profiles and Lifecycle Bias in Intergenerational Mobility Estimation*, J. Human Resources, vol. 51 no. 1 239-268.

OECD (2008), *Growing Unequal*, Paris.

OECD (2011), *Divided we Stand. Why Inequality Keeps Rising*, Paris.

OECD (2018), *A Broken Social Elevator? How to Promote Social Mobility*, OECD Publishing, Paris.

Salverda W., Nolan B., Checchi D., Marx I., McKnight A., Tóth I., van de Werfhorst H. (eds) (2014) *Changing Inequalities in Rich Countries. Analytical and Comparative Perspectives*, Oxford University Press.

Shorrocks, A. F., (1978b). *The measurement of mobility*, Econometrica Vol. 46, No. 5, pp. 1013-1024.

Shorrocks, A.F. (1980). *The Class of Additively Decomposable Inequality Measures,* Econometrica Vol. 48, No. 3, pp. 613-625.

Shorrocks, A.F. (1984). *Inequality decomposition by population subgroups,* Econometrica 52: 1369-85.

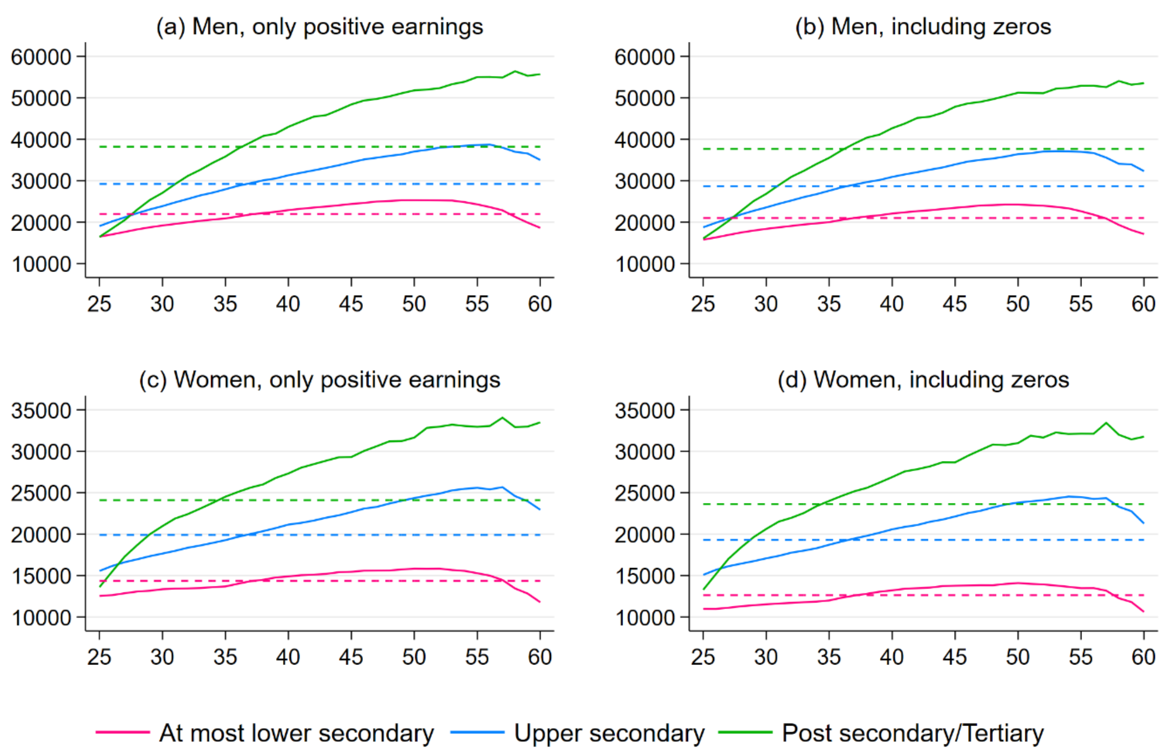
Struffolino E, Raitano M. (2020), *Labour market deregulation and employment trajectories: The increase in early-career complexity across cohorts in Italy*, *Social Indicators Research*, 151: 231-237.

APPENDIX A

We plot in Figure A1 the average real annual earnings against age, exploiting income records from AD-SILC for workers born between 1940 and 1974, divided by gender and highest level of education. We select a reasonable age window – from 25 to 60 years-old - in which every people willing to work should be in the labour market, after having completed formal education (or training) and before retirement.

We see that the average career pattern is increasing and concave for every level of education but, es expected, (i) there is level and growth premium for higher education, (ii) women's paths have a slowdown in the early-career stage which can be reasonably attributed to periods of maternity. Confirming this explanation, the timing of the slowdown for lower levels of education (around 25) precedes that of women with tertiary level education (around 30) who usually take more time to start a family after completing formal education. Moreover, when including zeros, the slowdown is more pronounced. Within this picture, what is relevant for our analysis is that, on average, current earnings coincide with lifetime earnings at some point between ages 35 to 40, regardless of gender and education level.

Figure A1: Average real annual earnings (€) by age



The mean career patterns are computed from AD-SILC for employees in the private sector born between 1940 and 1974, divided by gender and highest level of education attained.
The horizontal dashed lines represent mean individual earnings averaged over all workers.

APPENDIX B

We report below summary statistics of real annual earnings from labour for the balanced sample including only women.

Table B1: Summary statistics of real annual earnings (€) by cohort of birth

Cohort of birth	Period	N	Mean	SD	p10	p50	p90	Gini	GE2
1940-1944	1975-1989	8371	18960	8394	6679	19057	29449	0.246	0.098
1945-1949	1980-1994	11506	20936	9192	8525	20716	32634	0.242	0.096
1950-1954	1985-1999	12298	22027	9833	9527	21451	34359	0.244	0.100
1955-1959	1990-2004	14762	21815	10463	9096	20905	35459	0.263	0.115
1960-1964	1995-2009	19448	21148	11136	7932	20007	36533	0.290	0.139
1965-1969	2000-2014	26169	21525	11515	8474	20248	36939	0.291	0.143
1970-1974	2005-2018	21505	22182	11245	9553	20872	36917	0.275	0.128

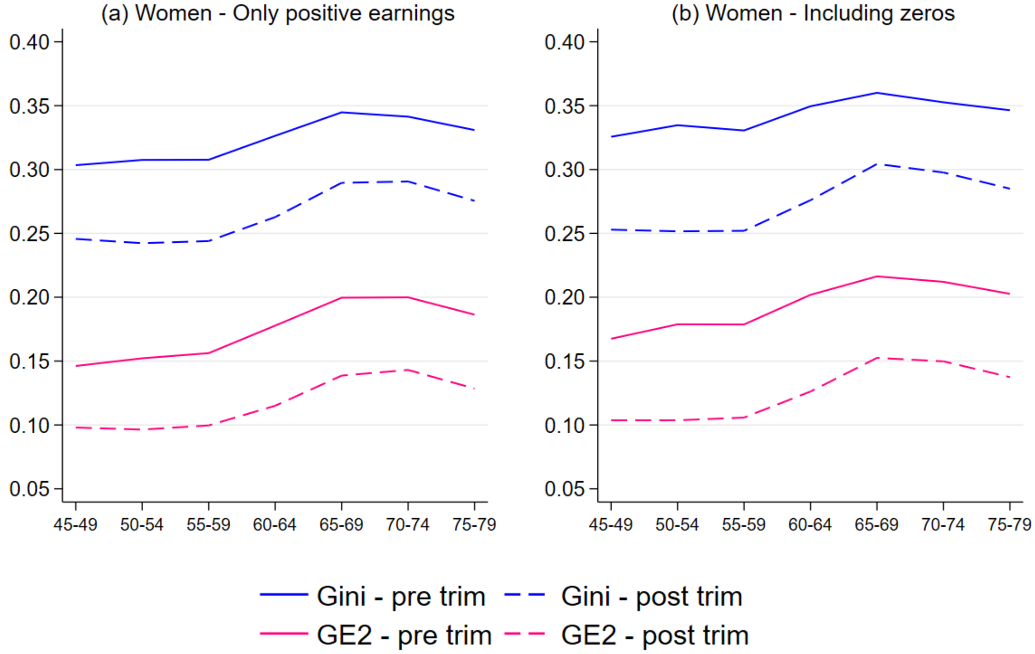
Table B2: Summary statistics of real annual earnings (€) by cohort of birth, including zeros

Cohort of birth	Period	N	Mean	SD	p10	p50	p90	Gini	GE2
1940-1944	1975-1989	8646	18710	8518	6003	18915	29311	0.253	0.104
1945-1949	1980-1994	12078	20608	9383	7690	20480	32522	0.252	0.104
1950-1954	1985-1999	12947	21669	9967	8876	21183	34182	0.252	0.106
1955-1959	1990-2004	15499	21289	10695	8346	20577	35160	0.276	0.126
1960-1964	1995-2009	20691	20479	11311	6905	19510	35965	0.304	0.153
1965-1969	2000-2014	27291	21162	11582	7866	19982	36471	0.298	0.150
1970-1974	2005-2018	23067	21702	11379	8879	20495	36471	0.285	0.137

The final female sample is composed of 10,369 (10,029) workers followed for 11 consecutive years in their mid-career excluding (including) zero earnings.

The figure below shows the impact of panel balancing for the women sample.

Figure B1: Impact of panel balancing, women only



The solid lines plot the inequality indices by cohort for workers aged from 35 to 46. The dashed lines plot the inequality indices by cohort computed excluding workers who are not present for 11 consecutive years in the sample.

Figure B2: Income risk decomposition by cohort of birth, women only

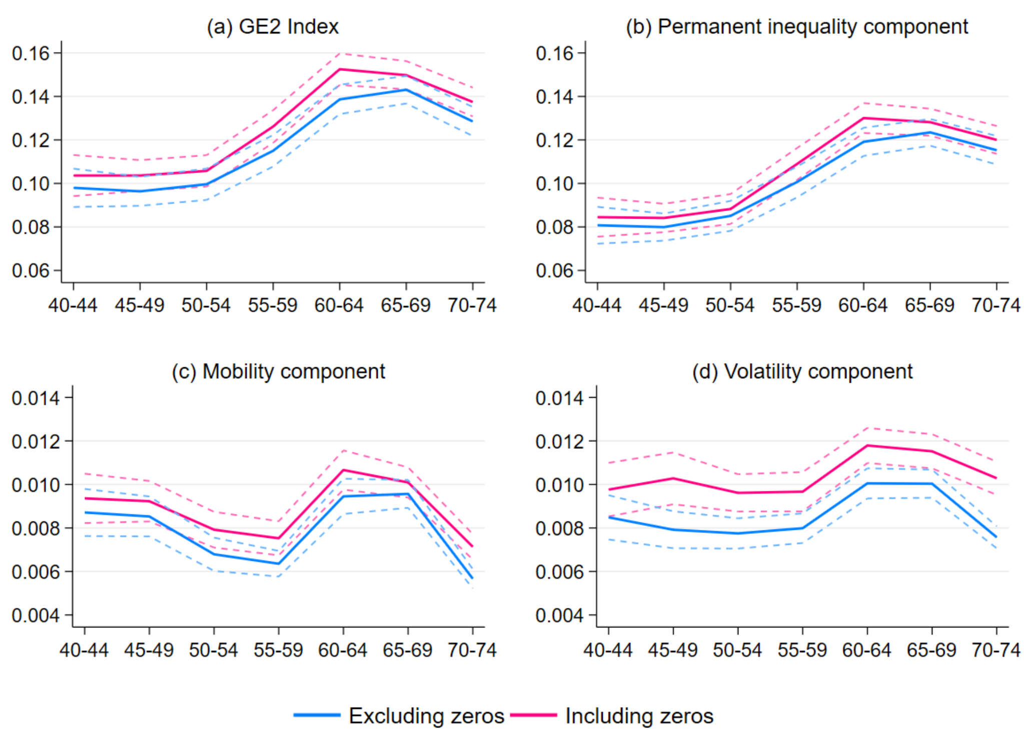
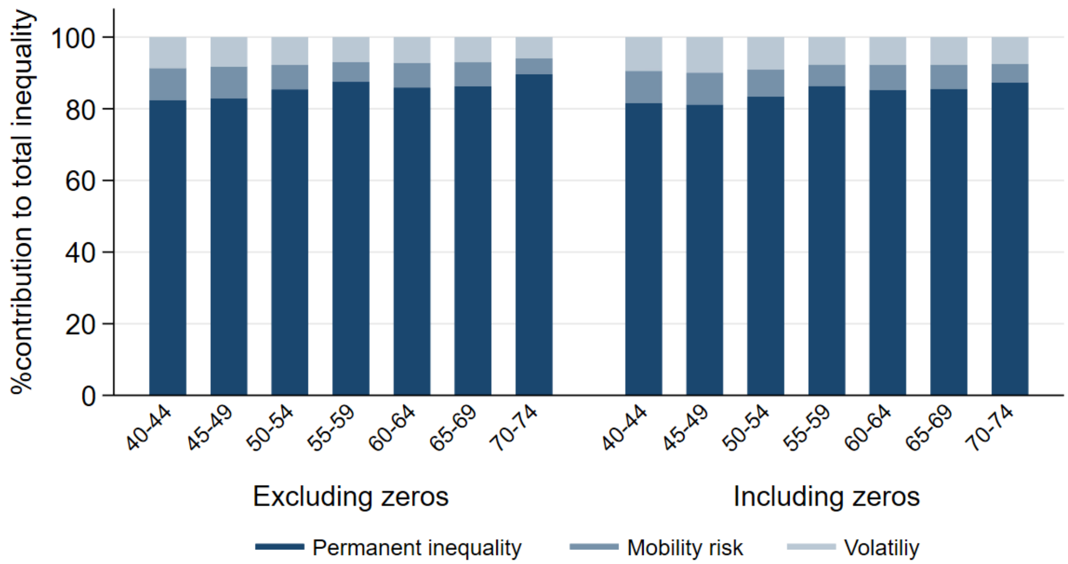


Figure 7: Contribution to overall inequality by cohort of birth, women only



APPENDIX C

Table C1: Income risk components, baseline sample

Table C2: % contribution to total inequality of income risk components, baseline sample

APPENDIX D

The table below reports the number of workers and observations for different age windows; the central age is 40 in all cases.

Table D1: Sample size for different age windows

	Baseline		Including zeros	
Window	**Workers**	**Observations**	**Workers**	**Observations**
T=5	22141	110705	21422	192798
T=7	21188	148316	22185	155295
T=9	20249	182241	23011	115055
T=11	18598	204578	19926	219186

We report below the results of the income risk decomposition for the baseline sample (male workers only, excluding zeros) using T=5, T=7 and T=9:

Figure D1: Income risk decomposition for male workers, by cohort of birth. T=5

Figure D2: Income risk decomposition for male workers, by cohort of birth. T=7

Figure D3: Income risk decomposition for male workers, by cohort of birth. T=9